THE CENTRE FOR MARKET AND PUBLIC ORGANISATION

# Student effort and educational attainment: Using the England football team to identify the education production function

Robert Metcalfe, Simon Burgess
and Steven Proud

November 2011

Working Paper No. 11/276

Centre for Market and Public Organisation
Bristol Institute of Public Affairs
University of Bristol
2 Priory Road
Bristol BS8 1TX
http://www.bristol.ac.uk/cmpo/

*Tel: (0117) 33 10952*
*Fax: (0117) 33 10705*
*E-mail: cmpo-admin@bristol.ac.uk*



ISSN 1473-625X

**CMPO Working Paper Series No. 11/276**

# Student effort and educational attainment: Using the England football team to identify the education production function

Robert Metcalfe[1], Simon Burgess[2]
and Steven Proud[3]

[1]*University of Oxford*
[2]*CMPO, University of Bristol*
[3]*University of Bristol*

November 2011

**Abstract**
We use a sharp, exogenous and repeated change in the value of leisure to identify the impact of student effort on educational achievement. The treatment arises from the partial overlap of the world's major international football tournaments with the exam period in England. Our data enable a clean difference-in-difference design. Performance is measured using the high-stakes tests that all students take at the end of compulsory schooling. We find a strongly significant effect: the average impact of a fall in effort is 0.12 SDs of student performance, significantly larger for male and disadvantaged students, as high as many educational policies.

**Keywords:** student effort, educational achievement, schools

**JEL Classification:** I20, J24

**Electronic version:** www.bristol.ac.uk/cmpo/publications/papers/2011/wp276.pdf

**Acknowledgements**
We are grateful to the DfE for access to the NPD data, and to the ESRC for funding through CMPO. We are grateful for comments to Rebecca Allen, Patty Anderson, Dan Hamermesh, Lindsey Macmillan, Andrew Oswald, Carol Propper, Helen Simpson, Sarah Smith, Liz Washbrook, Nicolas Ziebarth and seminar participants at ISER, the Geary Institute and CMPO.

**Address for correspondence**
CMPO, Bristol Institute of Public Affairs
University of Bristol
2 Priory Road
Bristol BS8 1TX
simon.burgess@bristol.ac.uk
www.bristol.ac.uk/cmpo/

## 1. Introduction

Understanding the education production function is a central goal of much research in education. Recent studies have analysed many components of this including the impact of class size, teacher quality, other school resources, peer groups, cognitive skills, family income, parental human capital and so on. The factor that we analyse here has received comparatively little attention, despite the fact that it is perhaps the one that the student herself focuses on most – her effort in studying is one of the few aspects of educational attainment actually under her own control. Does it matter? Does studying hard pay off? Data and identification problems are undoubtedly a major constraint: causal analysis of the impact of student effort on educational attainment requires an exogenous change in the marginal value of leisure or effort. This paper attempts to quantify how much student effort matters to educational achievement.

We use a sharp, exogenous and repeated change in the value of leisure to identify the impact of student effort on educational performance. Performance is measured using the universal high-stakes tests that students in schools in England take at the end of compulsory schooling. The treatment arises from the fact that the world's two most-watched international football tournaments (the FIFA World Cup and the UEFA European Championships) overlap with the exam period in schools in England, well known to be a nation obsessed with football. These tournaments are both attention-grabbing and highly salient for many students; they substantially raise the value of leisure time for many students and so are likely to reduce effort. They happen every other summer, so each year is sequentially either a treatment year or a control year. Because of the nature of the treatment and our data, we can implement a clean difference-in-difference design. We compare within-student variation in performance during the exam period between tournament and non-tournament years using seven years of student-subject data on practically all the students in England. This data allows us to bring out the heterogeneity of impact as well as quantifying the average effect.

In some ways, the treatment is ideal for a causal study. Exposure to the treatment is random: whether a particular student is born in an even year or an odd year. Neither students nor schools can affect the timing of the exams, which are always scheduled for the same period each year. The maximum potential treatment is very strong, as detailed below: the tournament always completely dominates TV, radio and other media during the weeks it takes place.  Because we do not observe actual hours spent thinking about the tournament, this is an intention to treat study, and actual treatment depends on an individual's interest in football; our data allow us to estimate the heterogeneity of effects. Another benefit of this dataset relative to other studies is that student effort is not confounded with teacher effort. There is no teaching at this time, leaving the student time to revise.

The key high-stakes examinations in England (called the General Certificate in Secondary Education or GCSE) are achievement tests, testing both knowledge and ability. They account for half of the final subject grade on average, and are always scheduled for May and June at the end of compulsory schooling (at age 16). We obtained data on exam timetables for each subject, and compare with the tournament dates. A proportion of exams overlap with these major football tournaments, and this generates within-student variation in tournament years.

We find a significant negative average effect of the tournament on exam performance, substantial for some groups. The mean deterioration in grades for subjects with exams during the tournament relative to earlier subjects is 0.063 standard deviations (SD)s of student performance. For highly affected groups such as male students[1], from disadvantaged families and of Black Caribbean heritage, the effect size is 0.10 SDs. Thus the overlap of exams and the tournaments reduces average attainment and raises educational inequality. In fact, since the summer exams only count for around half of the grade on average, the impact of the effort reduction directly on the exam score is about double these numbers. We consider the implications of these results for the role of effort in educational production in the Conclusion.

Our results contribute to the interpretation of two recent strands of work: the importance of non-cognitive and cognitive traits, and the effect of differences in school ethos on educational attainment. Firstly, a great deal of work has investigated the role of non-cognitive factors in educational attainment, alongside cognitive ability (see Cunha et al, 2010). Non-cognitive factors can be identified with personality traits (Heckman, 2011), and one of the 'big 5' personality traits is 'conscientiousness', with the related traits of self-control, accepting delayed gratification, and a strong work ethic (Heckman, 2011, p. 5); it is also related to the rate of time preference (Daly et al, 2009). Conscientiousness has been shown to be an excellent predictor of educational attainment and course grades (Almlund et al 2011, Borghans et al 2011). These aspects of self control and ability to concentrate are clearly related to the broad notion of effort we are using here. Our results add to this literature by isolating the effect of decisions on effort and time allocation in addition to the general ability to concentrate and exert self-control. The latter are differenced out in our design, leaving just the effect of differences in the contemporaneous value of leisure to influence achievement scores.

[1] An interest in football is in general more highly concentrated among males and lower socio-economic status groups in Europe than it is in the US.

Secondly, the work of Abdulkadiroglu et al (2011), Angrist et al (2010, 2011a, 2011b) and Dobbie and Fryer (2009) have focussed attention on “No Excuses” schools in the KIPP network or in the Harlem Children’s Zone. These schools all feature a long school day, a longer school year, very selective teacher recruitment, strong norms of behaviour, as well as other characteristics. Estimates of their impact on attainment, using lotteries as an identification strategy, show very powerful effects. This aggregate effect could be due to different aspects of the KIPP/HCZ ethos, but part of it is very likely to be increased effort from the students. Our results complement this by showing the impact of just a change in effort, and that that can have very substantial effects.

Other studies have addressed how effort matters for educational attainment. Stinebrickner & Stinebrickner (2008) found that randomly being assigned to a college roommate who has a video game console significantly reduces time allocated to studying (using self-completed surveys), which then negatively impacts on educational attainment. They use the roommate with the video console as an instrument to establish the causal impact of effort on attainment. The estimated effect may include peer effects as well as the changing marginal value of leisure. There is mixed evidence that class attendance at higher education is positively related to student performance[2]. These studies do not have clear identification strategies to demonstrate the marginal causal impact of changing the incentives of either effort or leisure. Using a birth cohort, de Fraja et al (2010) provide a theoretical and empirical model to test the impact of student, parent and teacher effort on attainment in the United Kingdom. Using the National Child Database Survey, they find that the students’ and parents’ effort are complements, and both seem more important to attainment than teacher effort. Effort and educational attainment are self-reported, and the former is based on a host of subjective variables about schoolwork, parents’ involvement in that schoolwork, and teacher involvement.

There is a small but growing literature within education on incentivising students to raise effort. Experiments in these studies increase the relative marginal value of effort over leisure. There have been some studies showing substantial positive effects of financial incentives on primary/elementary and secondary/high school students (Angrist et al, 2002; Henry & Rubinstein, 2002; Jackson, 2010; Kremer et al, 2009; Angrist & Lavy, 2009; Dearden et al, 2009; Dee, 2011; Pallais, 2009), although others demonstrate a lack of positive effects of financial incentives in educational attainment (Bettinger, 2010; Sharma, 2010; Fryer, 2010; and Rodriguez-Panas, 2010). Recent research from Levitt et al (2011a;b) show that financial incentives can increase attainment

[2] Some correlational studies found increased study time, and hence effort, to be associated with higher educational attainment (Schmidt, 1983; Michaels and Miethe, 1989; Douglas and Sulock, 1995; Young et al, 2003; and George et al 2008)), whereas others find no correlation (Park and Kerr, 1990; Nonis and Hudson, 2006) or even a negative ccorrelation (Krohn and O’Connor, 2005).

across all age categories, and very short-term incentives are particularly important in increasing examination performance.

More specifically, there is little evidence of how sporting events, that are not part of the traditional human capital production function, impact on attainment.[3] Other areas of life have been examined: Lozano (2008) considers the impact of the World Cup on the time use of adult male Americans.

The next section describes the English education system and our data. Section 3 sets out the framework underlying our approach, identification and selection issues. We present our results in section 4 and section 5 discusses the implications of our results.

## 2. Data

Compulsory education in England lasts for 11 years, in primary school from age 5 to 11, and secondary school from age 11 to 16. Teaching is organised around four Keystages, Key Stage 1 (KS1) up until age 7, KS2 to age 11, KS3 to age 14 and KS4 to age 16. There is essentially no grade repetition in England, each pupil moves up to the next school year each year. Each Keystage finishes with compulsory tests (though KS3 tests were recently discontinued). The KS4 exams at age 16, more widely known as GCSEs, are high stakes, crucial for continuing in school or looking for jobs. These exams are the focus of our empirical work.

Students take on average around eight subjects at GCSE, and most students will attempt at least five. Among these, English, maths and science are compulsory; others are optional chosen from a long list of possibilities, which will vary by school. These subjects are studied for two years up to the summer exams.

### a. Timing of Football Tournaments and Exams

Every four years (on even years) the FIFA World Cup takes place in June and July, and every other four years (on the different even years, so always two years apart) the UEFA European

[3] There is research that demonstrates that the impact of participation in sport is beneficial for educational attainment. There is some evidence from the United States that being involved in sport has a positive association with educational and labor market success (see for example Hanks, 1979; Long & Caudill, 1991; Maloney & McCormick, 1993; Barron et al., 2000; Eide & Ronan, 2001; Cornelißen & Pfeifer, 2010; Stevenson, 2010; Lechner, 2009). While there seems to be some positive impacts of participating in sport on education, the actual interest in sport on effort and educational attainment has received little attention.

Championships also take place in June and July.[4] The FIFA World Cup attracts a massive worldwide audience. For instance, the 2006 World Cup in Germany had television coverage in 214 countries around the world, with 73,000 hours of dedicated programming, which generated a total cumulative television audience of 26.29 billion people (FIFA, 2007). The UEFA European Championships are not as large as the World Cup, although the 2008 Euro tournament was watched live by at least 155 million TV viewers, and the final round of the tournament was shown in a total of 231 countries (UEFA, 2008).

Appendix Table 1 reports the time frame for the World Cups and European Championships from 2002 to 2008 (the years for which we have individual educational attainment scores in England). The English national football team qualified for the first three of these four international tournaments, but not for the 2008 European Championships[5]. We therefore classify 2008 as a “non-tournament” year, supported by the TV viewing data discussed below, but we test the robustness of our results to this decision.

In each of these years, the tournament overlapped with national UK examinations. We report the GCSE examination start dates and end dates in Appendix Table 2. There is no difference in exam dates between those years in which there is a football tournament and those years in which there is not. The proportion of exams during the football tournament ranges between 46% and 61%. The data on examination dates for each subject were obtained from Cambridge Examinations. Although different examination boards set their own exams, the exams of different boards for the same subject across the country are on the same day. The list of subjects used in our analysis is provided in Appendix Table 3. Some subjects have no exams during the tournament, others have a proportion of exams during the football period, and others have all the exams during the football period. Using the exact date of examination we calculate a variable that has the proportion of exams that are within the time period of the football tournament.

## b. Pupil Data

The data on pupils are taken from the National Pupil Database (NPD). The NPD is an administrative dataset of all pupils in the state-maintained system, some 93% of all pupils, made available to researchers by the Department for Education. It includes a census of pupils, taken each year in January, from 2002 onwards. In each cohort there is approximately half a million pupils, and so over

[4] The history and background to the FIFA World Cup and European Championships can be found at http://www.fifa.com/worldcup/ and http://www.uefa.com/ respectively.
[5] As some readers may know, England did not progress very far through the knock-out stages of any of these tournaments. We considered whether we could differentiate between exams sat before and after England were eliminated, but in fact the team did manage to remain in the tournament for almost all the exam period.

the seven year period we use, 2002 – 2008, we have some 3.5 million pupils. We have data on each student's gender, within-year age, ethnicity, an indicator of Special Educational Needs (SEN, which measures learning or behavioural difficulties), and eligibility for Free School Meals (FSM), which is dependent on eligibility for welfare benefits and is widely used as a measure of family poverty.

This census of personal characteristics can be linked to each pupil's test score history. Our analysis uses the subset of pupils that are identifiable within the state system throughout this period, which amounts to 90% of the cohort[6].

Our outcome variable is the pupil's performance in the high-stakes exams at the end of compulsory schooling at age 16, called GCSEs. These exams are nationally set and marked, and are marked remotely and independently from the school. Very few students resit exams, so there is almost always just the one grade per student per subject. The overall mark distribution is not normed each year, the average grade has been increasing year on year, which we control through common year effects. The structure of the attainment data is as follows: exams are nested within subjects which are nested within students. Students have grades for around 7 to 8 subjects[7] on average. Each of these subjects is graded in the same way, from A* to fail, and we assign numerical values to these letter grades using the National Curriculum points system. Each subject is assessed through a number of different instruments including typically a number of separate exams and coursework. The different exams for a subject are spread over a number of different days; on average, coursework contributes about half of the final grade, and we return to this when quantifying our results in section 4.e below.

We know each student's result in each subject, but the results for each individual exam are not available[8]. We know the date of each individual exam so we are able to characterise the degree of overlap between exams and the tournament for each subject. So while we cannot connect the date of a specific exam to the mark on a specific exam, we can relate the mark on a subject to the degree of overlap of the exams in that subject with the tournament.

We normalise the scores separately for each subject to remove any differences in subject difficulty; obviously the normalisation is done over all the years together as our focus is on across-year within-subject variation. Some analyses below use the student's overall mean exam score; for those analyses, we normalise the total score to a mean zero and SD of one. The SD of subject level scores

[6] Those that are excluded may have attended a private school for a period, may have spent time abroad (including Wales or Scotland), or may have been entirely educated in the English state system but their Unique Pupil Number was lost during a school transfer.
[7] We collapse Science entries to one subject.
[8] This rules out an even more fine-grained analysis, for example looking at the exam score the day after an important match.

and the SD of student averages over the subjects they take are about the same. The latter is only about 7.5% lower because a student's scores on all the subjects s/he takes are highly correlated.

As a measure of prior attainment, we use data on tests taken at age 11, Keystage 2 tests (KS2) in English, Maths and Science. These are compulsory for all students, and are also set and marked at a national level, remote from the school. One important and useful feature is that these tests are always taken in early May so are unaffected by the tournament taking place.

Table 1 provides an overview of our data. The basic facts are that around 12% of students are eligible for FSM, and around 85% of the students are white. For some analyses we take a subset of students who take both "late" and "early" exams, and we these account for 81.4% of our overall data. This subset differs a little: those students taking only "early" exams are slightly more likely to be poor, and have slightly lower prior attainment. Since the compulsory subjects have "early" exams, there are no students only taking "late" exams.

### c. Television Viewing Data

TV viewing data provide useful support for our assumption that watching the major international football tournaments is a very widespread phenomenon. The Broadcasters' Audience Research Board (www.barb.co.uk) provides weekly data on the viewing figures of the top 30 programmes per channel.

First, we show in Appendix table 4 how football programmes dominate the list of top programmes watched in this window. Football programmes are by far the most popular thing on television across the tournament period for all years, and more people watch BBC and ITV during football periods. Second we show in Figure 1 the big changes in TV viewing habits between June and April over our sample period. People generally watch much less TV in June, but in football tournament years, June viewing is considerably higher. This difference–in–difference in viewing habits supports our interpretation of the difference-in-difference in educational outcomes below, with peaks in 2002, 2004 and 2006. It also supports our assignment of 2008 as a non-tournament year. The observation for 2007 is rather out of line, but this seems to be due to an abnormally hot April that year reducing TV viewing then.

Watching television is the third most important activity (in terms of duration) that children over age 12 take part in (see Hofferth & Sandberg, 2001). In Europe, young people in lower SES families spend more time watching TV than higher SES families (Wartella, 1995; Hofferth & Sandberg, 2001), especially TV for entertainment (Hudson et al, 1999). Boys (aged 12 to 17) seem to watch more TV than similar aged girls (Robinson & Bianchi, 1997; Frederick, 1995).

## 3. Identification strategy

We set out a framework for the empirical work, to highlight the sources of identification in the different approaches we take below. We then discuss selection issues that might provide a threat to identification.

### a. Empirical Framework

The central idea is that the football tournaments dramatically raise the value of leisure time for some people, and correspondingly reduce the value of all other time uses. One of these time uses for students is the effort put into studying. Effort is to be understood in a broad sense as the number of hours spent preparing for the exams; this might include making the effort to ignore distractions and to create an environment to concentrate on study. We conjecture that the amount of productive study time is reduced both in the build-up to the tournament and then more significantly once the matches are under way.

A simple model of student effort is as follows. The level of educational attainment, as reflected in an exam result, is $q$. We assume that students value higher levels of attainment more highly because of the expected impact on future lifetime income, denoted $V(q)$. Students exert effort, $e$, when studying for their examinations, which has a utility cost, $\theta(e)$, increasing and convex in effort. Students have an underlying level of ability, a. Attainment is produced by ability and effort: $q = \delta(e, a)$. Student $i$ will choose $e_i$ to maximise $V\big(\delta(e_i, a_i)\big) - \theta(e_i)$. Optimal effort is the value of $e$ that solves: $V_i^{'}.\delta_i^{e}(e, a) = \theta_i^{'}(e)$, allowing all of the functions to vary by student.

Taking these components one by one, the effort function is central. The major cost of effort is the value of the leisure time forgone to undertake study. We assume that this may depend on observable individual characteristics such as gender, age, family background, denoted $Z_i$, and an unobservable individual component, $\mu_i$. The key factor for this paper is that the value of leisure increases for some individuals when a major football tournament takes place. We allow this impact of the tournament on the value of leisure to vary by individual, $\phi_i$. This taste for watching football will depend on cultural factors and an idiosyncratic component, which we expect to be substantial. The cultural factors may be associated with observable student characteristics, for example gender, ethnicity, social class, and location. An interest in football is by no means confined to men, but in England it remains a bigger part of male culture than female culture. Football has also been more associated with lower income and working class families (see Baker, 1979; Goldblatt, 2006). It may also matter whether a pupil lives in an urban or rural area since football is generally a city sport.

We denote the incidence of the tournament by the indicator terms *I(t = T)* and *I(m = T)*; that is: it is a tournament year, *t*, and that the tournament is actually taking place at the time of the relevant exam, *m*. We might expect an increase in the value of leisure in a tournament year even before the matches begin because of all the media attention and build up (denoted $\phi_i^0$), but we expect a larger impact once the tournament is actually under way (denoted $\phi_i^1$). So the overall cost of effort function for student *i* at time *t* is: $\theta_i^{'}(e) = \left(e; \mu_i, Z_i, \phi_i^0.\{I(t\ =\ T)\}, \phi_i^1.\{I(t\ =\ T).I(m\ =\ T)\}\right)$: effort depends on observable and unobservable student characteristics, and the individual impact of the tournament on the value of leisure, distinguishing between the general build-up and the actual coverage. This is a very flexible formulation.

The attainment technology, the impact of revision time and effort on the qualification gained, will also likely vary by student observable and unobservable characteristics, $\gamma_i$, and possibly by school, $s_i$. We allow for the possibility that the exam setting and marking may vary year-by-year by including year effects, *t*; these will net out of within-student estimates. We also allow student performance to vary through the exam period. There are many possibilities: for example, it may be that students learn and improve their exam technique as time goes by; or it could be that students tire and do worse on later exams; or it could be that later exams provide less time for last-minute revision; or it could be that, anticipating this, students over-revise for the later exams. In any case, we allow for unrestricted, idiosyncratic within-period time dummies, *m*. This is all summarised in the attainment function: $\delta_i^e(e, a) = (e;\ a_i, Z_i, \gamma_i, s_i, t, m)$. Conditional on what the student writes down in her exams, there is nothing in the setting or marking of the exams that could vary between tournament and non-tournament years beyond general time trends[9].

The student's valuation of the qualifications, $V_i^{'}$, may also depend on the same observable individual characteristics, $Z_i$, and an unobservable factor, $\varpi_i$.

Optimal effort chosen by the student depends on all these factors. Inserting these back into the attainment function gives the exam outcome. We allow for interactions so that the impact on the exam of an hour of study effort may depend on ability. Assuming a linear form we arrive at our empirical formulation:

[9] The exam marking occurs after the tournament is over.

$$q_{itm} = \beta_0 + \beta_1 . a_i + \beta_2 . Z_i + \sum_\tau \alpha_\tau . I(t = \tau) + \sum_n \pi_{in} . I(m = n) + f(\phi_i^0, a_i).\{I(t = T)\} + f(\phi_i^1, a_i).\{I(t = T).I(m = T)\} + \upsilon_i + \eta_{itm} \quad (1)$$

Where $\upsilon_i$ combines $\mu_i$, $\varpi_i$ and $\gamma_i$, plus any unmeasured aspects of ability (and also absorbs the school effect, $s_i$), and $\eta_{itm}$ is a noise term. In the results reported below, we also allow $f(\phi_i^1, a_i)$ to depend on $Z_i$.

In summary, the impact of the tournament on exam performance depends on the effect of the tournament on any reduction in study time (in turn depending on the student's taste for watching football), and the effectiveness of study time in raising exam scores (depending on the student's ability).

## b. Selection away from late exams

It would be potentially problematic if some students avoided taking optional subjects with late exams in tournament years. Even so, our conditional within-student difference should take out any first order effect of differences in unobservable student characteristics.

We compared the prior attainment and other characteristics of students picking late or early options in tournament and non-tournament years. We run a difference-in-difference, comparing the mean ability of students taking late options with those taking early options, and then difference that difference across T and NT years. Taking the student's average KS2 score as the measure of prior ability, the results show no effect. The diff-in-diff coefficient is positive, but only 0.001 of an SD and not significantly different from zero (even in a regression of 12.2m observations). So if the counter-story is that unobservably smarter pupils switch out of late options in anticipation of the tournament/exam clash some two and half years ahead, this is not supported by the evidence on observed ability.

In any case, there are reasons to believe this sort of selection not to be a serious issue. First, subjects are chosen over two years in advance of the exam period. While obviously the occurrence of the football tournaments is fully predictable, potential differential overlap of this with the exams is probably not a major reason for subject choice. Second, exam dates change, so even if students were attempting to strategise this they would be unable to do so perfectly. The exams varied between early and late for around half of all subjects over this period.

# 4. Results

We structure our results as follows. First, we present the simple aggregate time series story. Second, we look at simple differences between tournament and non-tournament years for matched groups of students. This gives an overall picture of the effect of tournaments but is vulnerable to being confounded with other year to year effects. So next we analyse within-individual differences and compare the distribution of these between tournament and non-tournament years. These provide our main results. This is an intention to treat study as we are estimating the effect of the potential treatment available. Fourth, we present robustness checks, and finally consider the effect size we have estimated relative to other education factors impacting on student performance.

Heterogeneity of response is likely to be very important in this context so in our results we emphasise the distribution of effects as much as the average. The effect we estimate depends on the change in the value of leisure time once the World Cup or European Championships is under way, and the impact of any reduction in study time on exam scores. The former at least is very likely to vary between students in a number of unobserved dimensions, and part of the role played by the observable characteristics of the students in the analysis is capturing a high or low interest in the tournament.

The most aggregate piece of evidence is the time series available on national average GCSE performance, which includes both state and private education. We focus on the percentage of students obtaining at least five good passes each year (grade C or higher) and Figure 2 plots the annual percentage change in this aggregate data from 1990 to 2008, with the tournament years highlighted. The visual impression that tournament years are associated with lower growth is supported by the means: a mean increase of 1.49 percentage points in tournament years against 1.63 in non-football years. However, this is just eighteen observations.

## a. Simple differences across years

We now exploit the pupil-level data, and start with simple differences: how students perform in tournament years against a similar set of students in non-tournament years. Each student is only present in one year, so we must compare groups of students across years. From our model, it is clear that a simple difference incorporates the pre-tournament build-up effect and the effect during the tournament itself. It also includes factors from the possibly-differing populations in tournament and non-tournament years and any differences in the general year dummies, so it is not cleanly identifying the effect of the tournament; that is in the following sub-section.

We first simply split our data up by gender, poverty status and ability (prior test performance), and tabulate for each of those groups the mean GCSE score in tournament-years minus the mean GCSE score in non-tournament-years, in student-level SD units[10]. The results are in Table 2. For example, taking high ability non-poor boys, we see that this group typically score 0.049 SDs lower in tournament years than non-tournament years. Across the table, about half of the entries are negative. The metric in this table is student-level SDs of GCSE scores. The magnitude of these effects is therefore not negligible. Comparing the column averages, there are negative averages for three of the four groups, larger for boys than girls. There is generally a greater difference for poor boys than non-poor boys. The overall row averages show a negative average for high ability students, and a positive effect for low ability students. But as is clear from our model, there are a number of other factors confounding the effect of the tournament in this analysis, and we turn to the difference-in-difference to identify the tournament effect itself.

## b. Difference-in-Differences across years and exam timing

A difference-in-difference analysis deals with this problem. Using our data on the timing of the exams and of the tournaments, we define 'late' subjects and 'early' subjects. In tournament years, late subjects are those in which at least two thirds of the exams are on dates overlapping the tournament, and early subjects are the rest. In non-tournament years, we take the same calendar dates in the tournament years to define late subjects. This allows us to compare performance in late and early exams within a year. As we noted above, it is likely that there will differences in performance on subjects late in the exam period versus early in the period for a number of reasons. We control for this and look for any differences in late - early performance gradients in tournament years.

For each pupil in each year, we define a late – early difference as the student's mean score over her/his late subjects minus her/his mean score over the early subjects. This is defined for the 82% of students with both late and early subjects. From (1), the expected late – early difference in a non-tournament year is simply: $\bar{q}_{i,t=NT,m=late} - \bar{q}_{i,t=NT,m=early} = \pi_{i,late} - \pi_{i,early}$, all observed and unobserved individual characteristics drop out, the year effect drops out leaving only that student's idiosyncratic performance change through the exam time. In tournament years, it is the same plus the impact of the tournament whilst it is in progress, $f\left(\phi_i^1, a_i\right)$. This permits a straightforward difference-in-difference to identify the effect of the tournament.

[10] We take out a common linear time trend, because GCSE results have trended up over our sample of years. This is the only analysis that we need to do this as all later analyses are based on within-student differences.

We compare the distribution of this gradient in tournament and non-tournament years in Figure 3. We see a wide range of outcomes, but that the weight of the distribution is above zero, that is, that somewhat more students perform better in late subjects than early ones. The figure also shows a difference between tournament and non-tournament years: the latter curve is shifted to the right, that is, the typical late-early gradient is lower in years when football tournaments are staged.

We use a regression analysis to summarise the differences in the distribution across tournament and non-tournament years. We run a regression of the gradient on observable characteristics, a dummy for whether it is a tournament year, and the interaction of the tournament dummy with all the observable characteristics, all with and without school fixed effects. We cluster standard errors at school level. The results are in Table 3.

Column 1 shows that the mean late-early difference is lower by 0.063 SDs in tournament years[11], and is unchanged by adding school fixed effects in column 3. Both these estimates are conditional on all our observable student characteristics (coefficients not reported). Both are very precisely estimated. We return to the quantitative significance below, but this is not trivial in magnitude.

Columns 2 and 4 introduce interactions of the observable student characteristics and the tournament indicator. Only a selection of interactions are reported in the Table and there are few other effects that are consistently significant across columns 2 and 4: students of Asian ethnicity other than Indian, Pakistani, Bangladeshi or Chinese have a significant negative effect; students of mixed white and Black Caribbean ethnicity a significant positive effect; students with special needs a consistent negative effect. The strongest single interaction is with gender: the impact of the tournament is greater in absolute terms by 0.025 SDs for males. This result ties in with the evidence noted above on relative TV viewing by gender and the much greater interest in football of boys than girls in England. Note, however, that the effect is statistically and quantitatively significant for female students too. Students from poor families also experience an additional negative impact of the tournament, adding 0.019 SDs in absolute terms to the base effect.

These effects are plausibly interpreted as mainly picking up cultural differences, as proxying a strong interest in watching the football tournament. But as we noted the impact of the tournament on exam scores in our model is the product of fewer hours of study and the effectiveness of those lost

[11] The tournament dummy is a linear combination of year dummies. Simply including a tournament dummy and omitting two year dummies obviously means that the reported tournament effect depends on which years are omitted. So to avoid an ad hoc choice, we decided to include a full set of dummies (omitting the constant) and compute the tournament effect as: ((d2002+d2004+d2006)/3) – (d2003+d2005+d2007+d2008)/4) and test whether this is different from zero. This is the standard error and significance level reported in the table. Note that the dependent variable is within-student late-early variation, so this is not capturing general drifts in average marks. We have also re-run the analysis with no year dummies at all and get similar results.

hours. The latter will depend on the student's ability, and the table shows results that confirm this view. Conditional on the other interactions, high ability students tend to suffer a higher penalty than the base effect, and low ability students less so. The extra hours of study lost due to the tournament are particularly costly to high ability students. Of course, ability is also likely to be correlated with cultural factors, but these effects appear to be swamped by the effectiveness factor.

We can display the differences between groups less parametrically by simply calculating the mean difference-in-difference for each of the groups we defined in Table 2. Table 4 reports the mean with the results of this calculation for the same simple twelve sub-groups as in Table 2. This is the average over all students in the group of their score in late exams in tournament years minus the mean score in early exams in tournament years, relative to the same calculation for non-tournament years. All the numbers are negative in the Table, and all are precisely defined and significantly different from zero. The size of the effects match up very well with the regression results in Table 3. The patterns across groups show bigger effects for boys than girls, bigger effects for disadvantaged groups, and a u-shape in ability, with the least effect in the middle group.

## c. Difference-in-Differences across years and exam timing within matched school-groups

We can exploit the large dataset to allow more of the response heterogeneity to be examined. We match groups at a very fine level, exploiting our very large and quite rich data. We match within school (around 2000 schools), and using the key observables of student gender, FSM status, prior attainment group[12], ethnic group[13] and quarter of birth. So each student in a tournament year is matched with a student in a non-tournament year in the same school and defined by the same set of observables. We take the mean (late exam score – early exam score) difference for each student, and then average this within each school*observables group, separately for tournament and non-tournament years, and analyse the difference.

Under our model (1), the within-individual difference removes all observable and unobservable individual characteristics, the year dummies and the effect of the general tournament build-up. The second difference across matched school-groups yields their mean direct tournament impact, $\overline{f(\phi^1, a)}$, plus any residual mean differences in unobservables across tournament years within our matched school*groups. The only threat to identification would be if, for example, the mean unobservable characteristics for poor, white, middle ability boys born in the first quarter of the year

[12] We use three broad groups, working below the expected level (Keystage 2 score below 27), working at the expected level (KS2 of 27), or above the expected level.
[13] We use four aggregated groups: White, Black, Asian and Other.

and attending a specific school differed significantly between the years (2002, 2004, 2006) and the years (2003, 2005, 2007, 2008), and differed in a way correlated with tournament years across the 400,000 school*groups.

To illustrate the heterogeneity of the impact Table 5 shows selected quantiles from the distribution of this difference-in-difference across some 15,000 school-groups with at least 40 students in. The overall median effect is 0.049 SDs. This is different to the equivalent regression value in Table 3 because that comes from an individual-level regression and the median is not weighted by school-group size; the weighted mean is -0.054.

But the focus of this table is the extremes not the average. The table reveals that some experience a very substantial effect. For example, 10% of groups have an effect more negative than -0.26 SDs. The 10th percentile for male students is -0.28 SDs and for disadvantaged students is -0.35 SDs.

We establish the statistical significance of the tournament effect within these matched school-groups as follows. Where *i* denotes student, *s* school, and *g* group, we run the following regression separately within each group: $\left(\bar{q}_{isg,late} - \bar{q}_{isg,early}\right) = a_g + b_g.(year = T) + \mu_{g,s(i)} + e_{isg}.$ And focus on the coefficients $b_g$. The within-student difference removes individual factors, the school fixed effects mean we are working off within-school differences between T and NT years, and the separate regressions by group mean we are comparing similar students.

The value of *b* for each of the 192 groups is plotted, ranked, in Figure 4, with standard error bars. The Figure shows that the great majority of groups see a reduction in the late-early performance gradient in tournament years. About half of the groups see a decline greater than 0.05 SDs and an important number of groups experience impacts greater than 0.1 SDs. A much smaller number of groups see positive impacts, but none of these are significant. The size of the standard errors depends on group size, but of the 192 groups, the coefficient is significantly negative (at 5%) in 104 and significantly positive in none. The 104 groups are the largest groups and cover 91% of students.

The Figure displays in sub-panels the variation within gender groups, within poverty status, ability and ethnicity groups[14]. The groups ranked within gender show that throughout the range, the *x*-ranked make group is more affected than the *x*-ranked female group, with a roughly constant gap between them. The great majority of both groups see a negative effect. The sub-panel splitting by poverty status shows that the most affected poor groups suffer a greater impact than the most affected non-poor groups, though for much of the range, the two groups experience broadly similar levels of impact. In the sub-panel by ability, the three lines are generally not far apart. What

[14] We take out the otherwise-equivalent groups to those missing for the other subsets – so for example, if poor*females*high ability*white*first qtr birth is missing, we delete the same non-poor group.

differences there are suggest that the middle ability group experiences the least impact among all but the most affected groups. The high ability groups see the greatest impact throughout almost all the distribution.

## d. Robustness checks

We review the sensitivity of our results to a number of the data issues we have dealt with. First, instead of defining "late" subjects as those in which two thirds of exams overlapped with the tournament, we used a half as the cut-off. The average effect is still strongly negative, but as this definition is less sharp, employing this produces a lower estimated effect of the tournament: an overall average of -0.038 in the equivalent of Table 4 rather than -0.056. There is also less of a distinction between the genders.

Second, it is possible in the English system to sit exams a year earlier than the age-correct year. This is not uncommon in maths, and fairly rare in other subjects. If we drop students who sit maths early, the point estimate is unchanged at -0.056.

Third, we counted the year 2008 as a non-tournament year for students in England as the national team did not qualify for the European Championships. This decision was based in part on the TV viewing figures in Figure 1. If instead we designate it as a tournament year, we estimate a reduced negative effect for boys, and a positive effect for girls. It is unfortunate that the ambiguous status of this year coincides with it being a year when one of the compulsory subjects, science, switches from being 'early' to being 'late', giving that year some leverage. If we drop all science results from the estimation and compare taking 2008 as a tournament and non-tournament year, the estimation of the tournament effect barely changes.

Fourth, we can extend the simple dichotomous early/late subject variable and construct a continuous variable from the exam timetable information. The variable $P_{itz}$ is the proportion of examinations in subject *z* that are taken during the tournament period by student *i* in year *t*. We modify our main model (1) as follows, indexed now by subject *z* rather than by exam-month *m*, and collecting all the observed and unobserved student characteristics in $S_i$:

$$q_{itz} = \beta_0 + \beta_1 . S_i + \alpha_{t(i)} . t(i) + \phi_i . I(t = F) . P_{itz} + \varphi_i . P_{itz} + \zeta_{itz} \qquad (2)$$

Note that each student only appears in one year and so only one time dummy is relevant. The term $\varphi_i$ represents the early – late gradient in exam scores in non-tournament years and $(\varphi_i + \phi_i)$ is the gradient in tournament years.

The within-estimator simply relates the exam scores across subjects within-student with the timetabling of those exams[15]. The source of variation is simply how many of the exams were timetabled later in the period, and differencing the impact of that timing across tournament and non-tournament years. We also estimate a pooled version of (2) including observable student characteristics and no student fixed effects. In this, there is again variation across subjects and across years for a given subject.

The results are in Table 6, the three columns reporting the two key coefficients of interest, $\bar{\varphi}$ and $\bar{\phi}$, for models with (1) just year dummies, (2) year dummies plus individual characteristics, and (3) year dummies and student fixed effects. As the residual error term, $\zeta_{itz}$ is likely to be correlated across each pupils' exam results, we cluster standard deviations at the individual level. The results suggest that typically students do better in later exams than earlier ones. This difference is significantly less in tournament years, the coefficient $\bar{\phi}$ being negative and precisely determined, confirming our main findings.

## e. Quantifying the effect

Our main results are the difference in difference analysis in the later tables. These are cleanly identified and seem unlikely to be confounded, but the late-early difference may not be the full story if the early results are affected by the build-up to the tournament. The results in Table 2 give the full effect but are likely to be confounded by other common year effects more complex than a simple trend. We illustrate how the two different results relate to one another in Figure 5. This shows the difference in the late - early difference between tournament and non-tournament years (from Tables 3, 4 and 5), and the overall effect; in principle, either one could be larger depending on the degree of impact of the pre-tournament build up, and the fraction of late exams. We focus on the difference-in-difference results.

We present the average effect and a representative high effect. The latter serves two purposes: it shows how affected the most affected groups are, and it allows us to address the importance of effort in general by asking how much do scores fall when students significantly reduce their effort. We present the results in terms of effect sizes (standard deviation units for student average test scores) and in terms of GCSE grades (counting the difference between A and B as one) for local policy interest.

Our results give the impact of the tournament on the overall GCSE score, and this is the appropriate outcome for addressing the question of whether the exam timetable should be shifted. But for the

[15] Recall that potential differences in the difficulty of the subject are dealt with by normalising the scores by subject.

broader issue of the role of effort in the educational production function we need to focus on the impact of the tournament on performance in the achievement tests. And to do this we need to take account of the fact that the summer exams only form part of the overall grade, the rest being made up of coursework and continual assessment. For instance, in English, coursework can typically account for 40% of the overall GCSE grade. The fraction of the mark from the final summer exams varies by course and by year[16]. Assembling information from a range of sources:

Year 11 summer exam contribution

| Subject | % | Subject | % | Subject | % |
|---|---|---|---|---|---|
| | | | | | |
| English Language | 60 | English Literature | 70 | Maths | 55 – 70 |
| Science | 50 | Child Development | 50 | IT | 40 – 50 |
| Languages | 50 | Geography | 75 | History | 75 |
| Design & technology | 40 | Law | 80 | Media Studies | 50 |
| Music | 25 | Physical Ed'n (gym) | 50 | Religious Education | 75 – 100 |

The three compulsory subjects average out to about 60% of the final mark coming from the summer exams, and the other subjects not that different. In this case, the direct impact of a reduction in student effort on exam performance is about double the regression coefficients, and this is what we report below.

The results are in Table 7; column 1 doubles the coefficients from the tables specified. To get to the numbers in column 2, note that the dependent variable is {(pupil mean score in late exams) – (pupil mean score in early exams)} measured in subject-SD units. The dummy on "tournament year" means that the mean late exam score is lower by the coefficient in those years. The effect on the overall average is this number times the fraction of exams that are scheduled late, which is (1.75/7.80); we also convert from subject-SD units to pupil-SD units. The effect of effort on late exams is necessarily larger than the impact of the tournament on the overall student average grade, but the latter is non-trivial.

To put these numbers in perspective, we can compare to a range of policy effects (see Jacob and Ludwig, 2008, and Dobbie and Fryer, 2009). For example, the "Literacy Hour" intervention in England raised reading attainment by 0.06 SDs (Machin & McNally, 2008). A unit SD increase in teacher quality raises test scores by around 0.15 to 0.24 SDs per year, 0.27 in England (Rockoff, 2004; Rivkin, Hanushek and Kain, 2005; Aaronson, Barrow, and Sander, 2007; Kane and Staiger, 2008; Slater, Davies and Burgess, 2011). The effect of major "early years" programmes such as Head Start is 0.147

[16] From 2009, coursework has been phased out in favour of "controlled assessment", and the system has changed.

SDs in applied problems and 0.319 in letter identification (Currie and Thomas, 1995; Ludwig and Phillips, 2007). Crawford et al (2007) have shown that a student's month of birth has effects on GCSE outcomes: students who have spent their entire school careers as the youngest in the class (August-borns) score on average 0.116 SDs (girls) or 0.131 SDs (boys) lower than the oldest in the class (September-born students). Substantial effects on pupil progress have been found in "No Excuses" Charter schools, of between 0.10 - 0.40 standard deviations increase per year in mathematics and reading (Abdulkadiroglu et al, 2009; Angrist et al, 2010). More closely related to our focus on effort, Fryer (2010) and Levitt et al (2011a) show that incentivising students to raise their effort (input-based student incentives) have an effect size of about 0.15 SDs, and Levitt et al (2011b) show that incentives on the day of a test can increase test scores by around 0.2 SDs.

In this context, our results suggest that effort clearly has a substantial impact on student performance, and can have a similar effect to raising teacher quality by one SD.

## 5. Conclusion

We use a sharp, exogenous and repeated change in the value of leisure to identify the impact of student effort on educational performance. This arises from the fact that the world's most watched international football tournaments overlap with the main high-stakes exam period in schools in England, well known to be a nation obsessed with football. We compare within-student variation in exam performance during the exam period between tournament and non-tournament years using seven years of student-subject data on practically all the students in England. This data allows us to bring out the heterogeneity of impact as well as quantifying the average effect.

We find a significant negative average effect of the tournament on exam performance, substantial for particular groups. The mean deterioration in grades for subjects with exams during the tournament relative to earlier subjects is 0.063 SDs of student performance. Since exams only account for half of the grade on average, this means that the impact of the tournament on just the exams is double that at 0.12 SDs. For highly affected groups such as low ability male students from disadvantaged families, this impact is over 0.2SDs. In the context of the range of effect sizes reported by Jacob and Ludwig (2008), these are non-trivial amounts.

We think of this impact arising through a reduction in student effort, with that time being spent instead on watching the football tournament. The variation in impact arises because of differing tastes for football, arising in turn from cultural norms and idiosyncratic factors, and from the differential effectiveness of an hour of study on exam performance. There are a number of contributory factors, worthy of further research. Firstly, the allocation of time by children is a largely ignored topic by economists (Larson and Verma, 1999), but could end up being crucial to educational attainment. Understanding this, along with their rate of time preference, and whether they are maximising their utility by watching the football as opposed to revising for their examinations, will be important for unpicking this relationship. Secondly, it could be that students' parents are watching the football tournament and pay less attention to their child revising for their examinations, which lowers their encouragement or confidence to the child during their examinations (Barnard, 2004).

Our results relate to two issues: a broad debate on the nature of educational attainment function, and a policy question specific to England about the timing of summer exams.

Taking the second first, our results show that having important exams in the tournament period reduces educational attainment at the median by 0.02 SDs of pupil average scores. We estimate much greater negative effects for male students, students from disadvantaged families and low ability students. These are groups that have lower performance anyway, and our results show that the tournament has a substantial negative effect on their performance. This in turn will affect their likelihood of progression through the educational system and their lifetime income. Given this, the benefits of moving the exams just a few weeks earlier in tournament years are significant. By comparison, the costs are likely to be transient. More generally, scheduling GCSE examinations during football tournaments lowers overall human capital in the UK, with implications for future economic growth.

Our results carry a number of implications for our understanding of the educational production function. The first is simply to note that student effort matters a lot. The coefficients directly show the impact on subject grades for late subjects. For strongly affected groups, this is as high as 0.1 SDs. As we noted above, this is the impact on the overall GCSE score, and given that exam performance is typically worth about half of the overall score, the impact of reduced effort in the exam period on the exam score will be roughly twice that, 0.2SDs. This is a very substantial effect. Obviously, it would be wrong to extrapolate from this number to a longer-term reduction in effort, as we have captured the effect of a reduction in high-value effort just before the exam. Nevertheless, comparing two otherwise identical students of average ability, the one putting in considerably less effort to their school work will perform substantially less well, at least 0.2 SDs worse, and conceivably worse still.

This matters for a number of reasons. First, unlike genetic characteristics, cognitive ability or non-cognitive traits, effort is almost immediately changeable. Our results suggest that this could have a big effect. This ties in with recent results on policies aimed at raising attainment. Fryer's (2010), Jackson's (2010) and Levitt et al's (2011a;b) results suggest that directly paying students for greater effort has an impact on test scores. Furthermore, the dramatic test score gains cited for "No Excuses" schools in the KIPP and HCZ or some Charter schools (Abdulkadiroglu et al. (2011), Angrist et al.(2011), Fryer and Dobbie (2009)) can plausibly be interpreted as those environments eliciting greater effort from the students. The fact that we find changes in student effort to be very potent in affecting test scores suggests that policy levers to raise effort through incentives or changing school ethos are worth considering seriously. Such interventions would be justified if the low effort resulted from market failures due to lack of information on the returns to schooling, or time-inconsistent discounting.

Secondly, the importance of a manipulable factor such as effort for adolescents' educational performance provides evidence of potentially high value policy interventions much later than "early years" policies. This is encouraging, offering some hope that low performing students' trajectories in life can perhaps be effectively improved even after a difficult environment early in life.

Finally, there are suggestions from neuroscience that activation of the brain's motivational circuitry directly affects cognitive learning processes involving many different regions throughout the brain (Adcock,2006; Bavelier et al. (2010); Howard-Jones (2011). If so, this means that students devoting a lot of time to their studies are likely to have long-lasting effects on their ability in addition to the immediate effects on their test scores.

## References


Aaronson, D., Barrow, L. and Sander, W. (2007) Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics*, 25: 1, pp. 95-135.

Abdulkadiroglu, A., Angrist, J., Cohodes, S., Dynarski, S., Fullerton, J., Kane, T. and Pathak, P. (2011) "Informing the Debate: Comparing Boston's Charter, Pilot and Traditional Schools." Forthcoming, *Quarterly Journal of Economics*

Adcock, R. A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B. and Gabrieli, J. D. E. (2006) Reward-motivated learning: Mesolimbic activation precedes memory formation. *Neuron*, 50: 3, pp. 507-517.

Almlund, M., A. Duckworth, J. J. Heckman, and T. Kautz (2011). Personality psychology and economics. In E. A. Hanushek, S. Machin, and L. Woessmann (Eds.), Handbook of the Economics of Education, Volume 4. Amsterdam: Elsevier. Forthcoming.

Angrist, J. and Lavy, V. (2009) The Effects of High Stakes High School Achievement Awards: Evidence from a Randomized Trial. *American Economic Review*, 99: 4, pp. 1384-1414.

Angrist, J. D. Dynarski, S.M., Kane, T.J., Pathak, P. A. and Walters, C. (2010) Inputs and Impacts in Charter Schools: KIPP Lynn. *American Economic Review Papers and Proceedings* 100 (2).

Angrist, A., Dynarski, S., Kane, T., Pathak, P, and Walters, C. (2011) Who benefits from KIPP? IZA DP 5690

Angrist, A., Pathak, P, and Walters, C. (2011) Explaining Charter School Effectiveness. NBER WP 17332.

Angrist, J., Bettinger, E., Bloom, E., King, E. and Kremer, M. (2002) Vouchers for private schooling in Colombia: Evidence from a randomized natural experiment. *American Economic Review*, 92: 5, pp. 1535-1558.

Angrist, J., Oreopoulos, P. and Williams, T. (2010) When Opportunity Knocks, Who Answers? New Evidence on College Achievement Awards *NBER working paper series no. w16643* (Cambridge, Mass., National Bureau of Economic Research).

Baker, W. J. (1979) The Making of a Working-Class Football Culture in Victorian England. *Journal of Social History*, 13: 2, pp. 241-251.

Barnard, W. M. (2004) Parent involvement in elementary school and educational attainment. *Children and Youth Services Review*, 26: 1, pp. 39-62.

Barron, J. M., Ewing, B. T. and Waddell, G. R. (2000) The effects of high school athletic participation on education and labor market outcomes. *Review of Economics and Statistics*, 82: 3, pp. 409-421.

Bavelier, D., Levi, D. M., Li, R. W., Dan, Y. and Hensch, T. K. (2010) Removing Brakes on Adult Brain Plasticity: From Molecular to Behavioral Interventions. *Journal of Neuroscience*, 30: 45, pp. 14964-14971.

Bettinger, E. P. (2010) Paying to Learn The Effect of Financial Incentives on Elementary School Test

Scores *NBER working paper series no. w16333* (Cambridge, Mass., National Bureau of Economic Research).

Borghans, L., B. H. H. Golsteyn, J. J. Heckman, and J. E. Humphries (2011). IQ, achievement, and personality. Unpublished manuscript, University of Maastricht and University of Chicago (revised from the 2009 version).

Clark, D. B. (2004) The natural history of adolescent alcohol use disorders. *Addiction*, 99, pp. 5-22.

Crawford, C., Dearden, L. and Meghir, C. (2007) When You Are Born Matters: The Impact of Date of Birth on Child Cognitive Outcomes in England, Centre for the Economics of Education, LSE).

Cunha, F., J. J. Heckman, and S. M. Schennach (2010, May). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica* 78 (3), 883-931.

Currie, J. and Thomas, D. (1995) "Does Head Start Make a Difference?" American Economic Review 85(3), 341–364

Daly, M., L. Delaney, and C. P. Harmon (2009). Psychological and biological foundations of time preferences. *Journal of the European Economic Association* 7 (2-3), 659-669.

De Fraja, G., Oliveira, T. and Zanchi, L. (2010) Must Try Harder: Evaluating the Role of Effort in Educational Attainment. *Review of Economics and Statistics*, 92: 3, pp. 577-597.

Dearden, L., Emmerson, C., Frayne, C. and Meghir, C. (2009) Conditional Cash Transfers and School Dropout Rates. *Journal of Human Resources*, 44: 4, pp. 827-857.

Dee, T. S. (2011) Conditional cash penalties in education: Evidence from the Learnfare experiment. *Economics of Education Review*, 30: 5, pp. 924-937.

Dobbie, W. and Fryer Jr., R. G. (2009) Are high quality schools enough to close the achievement gap? Evidence from a social experiment in Harlem. NBER WP 15473

Douglas, S. and Sulock, J. (1995) Estimating Educational Production-Functions with Correction for Drops. *Journal of Economic Education*, 26: 2, pp. 101-112.

Eide, E. R. and Ronan, N. (2001) Is participation in high school athletics an investment or a consumption good? Evidence from high school and beyond. *Economics of Education Review*, 20: 5, pp. 431-442.

Federation International de Football Association (FIFA),((2007)) FIFA World Cup and Television. http://www.fifa.com/mm/document/fifafacts/ffprojects/ip-401_06e_tv_2658.pdf

Frederick, J.A. (1995). As time goes by. Time use of Canadians: General social survey. Statistics Canada: Ontario.

Fryer, R. G., Jr. (2010) Financial Incentives and Student Achievement Evidence from Randomized Trials NBER WP 15898

George, D., Dixon, S., Stansal, E., Gelb, S. L. and Pheri, T. (2008) Time diary and questionnaire assessment of factors associated with academic and personal success among university undergraduates. *Journal of American College Health*, 56: 6, pp. 706-715.

Goldblatt, D. (2006) *The ball is round: a global history of football*. (London, Viking).

Hanks, M. (1979) Race, Sexual Status and Athletics in the Process of Educational-Achievement. *Social Science Quarterly*, 60: 3, pp. 482-496.

Hayward-Jones, P. A. (2011) The impact of digital technologies on human wellbeing, Oxford, Nominet Trust).

Heckman, J. J. (2011) Integrating personality psychology into economics. NBER WP 17378.

Henry, G. T. and Rubenstein, R. (2002) Paying for grades: Impact of merit-based financial aid on educational quality. *Journal of Policy Analysis and Management*, 21: 1, pp. 93-109.

Hofferth, Sandra L. & Sandberg, J.F. (2001). How American Children Spend Their Time. *Journal of Marriage and Family*, 63: 295-308.

Jackson, C. K. (2010) A Little Now for a Lot Later A Look at a Texas Advanced Placement Incentive Program. *Journal of Human Resources*, 45: 3, pp. 591-639.

Jacob, B., Ludwig, J. and National Bureau of Economic Research. (2008) Improving Educational Outcomes for Poor Children *NBER working paper series no. w14550* (Cambridge, Mass., National Bureau of Economic Research).

Kane, T. and Staiger, D. (2008) “Estimating Teacher Impacts on Student Achievement: An Experimental Validation.” NBER Working Paper 14607.

Kremer, M., Miguel, E. and Thornton, R. (2009) Incentives to Learn. *Review of Economics and Statistics*, 91: 3, pp. 437-456.

Krohn, G. A. and O'Connor, C. M. (2005) Student effort and performance over the semester. *Journal of Economic Education*, 36: 1, pp. 3-28.

Krueger, A. B. (1999) Experimental estimates of education production functions. *Quarterly Journal of Economics*, 114: 2, pp. 497-532.

Larson, R. W. and Verma, S. (1999) How children and adolescents spend time across the world: Work, play, and developmental opportunities. *Psychological Bulletin*, 125: 6, pp. 701-736.

Lechner, M. (2009) Long-run labour market and health effects of individual sports activities. *Journal of Health Economics*, 28: 4, pp. 839-854.

Levitt, S., List, J. and Sadoff, S. (2011a) The Effect of Performance-Based Incentives on Educational Achievement: Evidence from a Randomized Experiment, *mimeo*.

Levitt, S., List, J. and Sadoff, S. (2011a). The Effect of Performance-Based Incentives on Educational Achievement: Evidence from a Randomized Experiment. Mimeo, University of Chicago.

Levitt, S., List, J. and Sadoff, S. (2011b) The Impact of Short-Term Financial Incentives on Student Performance, *mimeo*.

Levitt, S., List, J. and Sadoff, S.. (2011b). The Impact of Short-Term Financial Incentives on Student Performance. Mimeo, University of Chicago.

Long, J. E. and Caudill, S. B. (1991) The Impact of Participation in Intercollegiate Athletics on Income and Graduation. *Review of Economics and Statistics*, 73: 3, pp. 525-531.

Lozano, F. (2008) The FIFA World Cup and the hours of work of American Males (Pomona College.

Ludwig, J. and Phillips, D. (2007) “The Benefits and Costs of Head Start.” NBER WorkingPaper 12973.

Machin, S. and McNally, S. (2008) The literacy hour. *Journal of Public Economics*, 92: 5-6, pp. 1441-1462.

Maloney, M. T. and Mccormick, R. E. (1993) An Examination of the Role That Intercollegiate Athletic Participation Plays in Academic-Achievement - Athletes Feats in the Classroom. *Journal of Human Resources*, 28: 3, pp. 555-570.

Michaels, J. W. and Miethe, T. D. (1989) Academic Effort and College Grades. *Social Forces*, 68: 1, pp. 309-319.

Nonis, S. A. and Hudson, G. I. (2006) Academic Performance of College Students: Influence of Time Spent Studying and Working. *Journal of Education for Business*, 81: 3, pp. 151-159.

Pallais, A. (2009) Taking a Chance on College Is the Tennessee Education Lottery Scholarship Program a Winner? *Journal of Human Resources*, 44: 1, pp. 199-222.

Park, K. H. and Kerr, P. M. (1990) Determinants of Academic-Performance - a Multinomial Logit Approach. *Journal of Economic Education*, 21: 2, pp. 101-111.

Renna, F. (2007) The economic cost of teen drinking: Late graduation and lowered earnings. *Health Economics*, 16: 4, pp. 407-419.

Rivkin, S. G., Hanushek, E. A. and Kain, J. F. (2005), Teachers, Schools, and Academic Achievement. Econometrica, 73: 417–458

Robinson, J. & Bianchi, S. (1997). The children’s hours. *American Demographics*, 19: 20-24.

Rockoff, J. E. (2004) “The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data.” American Economic Review, 94(2): 247–252.

Rodriguez-Planas, N. (2010) Longer-Term Impacts of Mentoring, Educational Services, and Incentives to Learn: Evidence from a Randomized Trial *IZA Discussion Papers 4754*, Institute for the Study of Labor (IZA)).

Schmidt, R. M. (1983) Who Maximizes What - a Study in Student Time Allocation. *American Economic Review*, 73: 2, pp. 23-28.

Sharma, D. (2010) The Impact of Financial Incentives on Academic Achievement and Household Behavior: Evidence from a Randomized Trial in Nepal *Working Paper*, The Ohio State University).

Slater, H., Davies, N. and Burgess, S. (2011) Do teachers matter? Measuring the variation in teacher effectiveness in England. Forthcoming Oxford Bulletin of Economics and Statistics

Stevenson, B. (2010) Beyond the Classroom: Using Title Ix to Measure the Return to High School Sports. *Review of Economics and Statistics*, 92: 2, pp. 284-301.

Stinebrickner, R. and Stinebrickner, T. R. (2008) The causal effect of studying on academic performance. *B E Journal of Economic Analysis & Policy*, 8: 1.

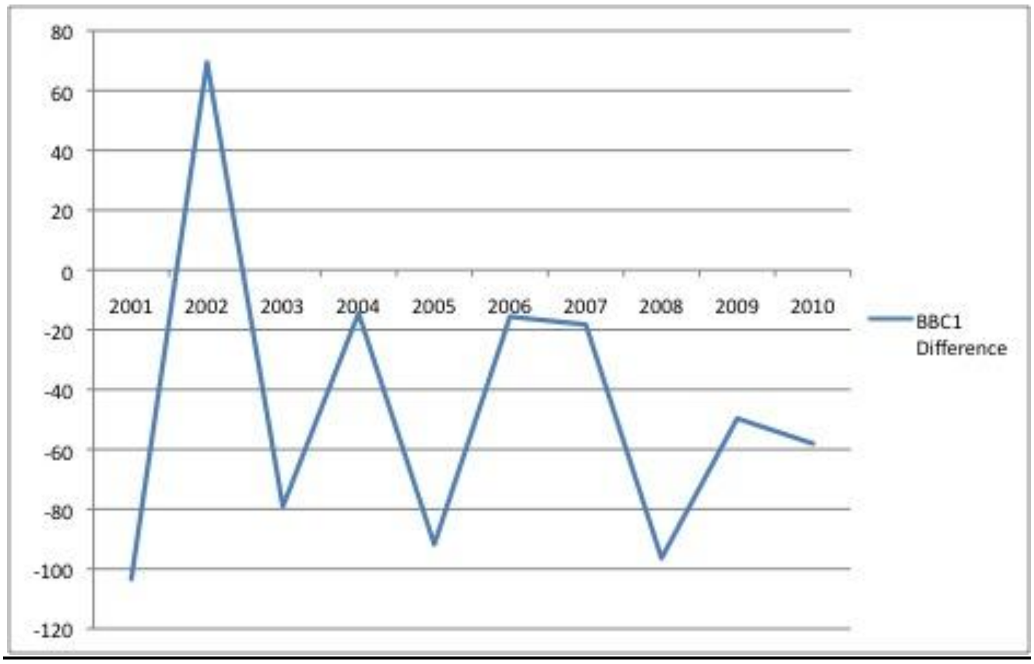
UEFA (2004) UEFA Euro 2008™ Review.

Wartella, E. (1995). Media and problem behaviors in young people. In Rutter, M. & Smith, D.J. (Eds.) Psychosocial disorders in young people: Time trends and their causes. Wiley: New York, pp. 296-323.

Young, M. R., Klemz, B. R. and Murphy, J. W. (2003) Enhancing Learning Outcomes: The Effects of Instructional Technology, Learning Styles, Instructional Methods, and Student Behavior. *Journal of Marketing Education*, 25: 2, pp. 130-142.

## Figure 1: Major Football Tournaments and TV viewing figures

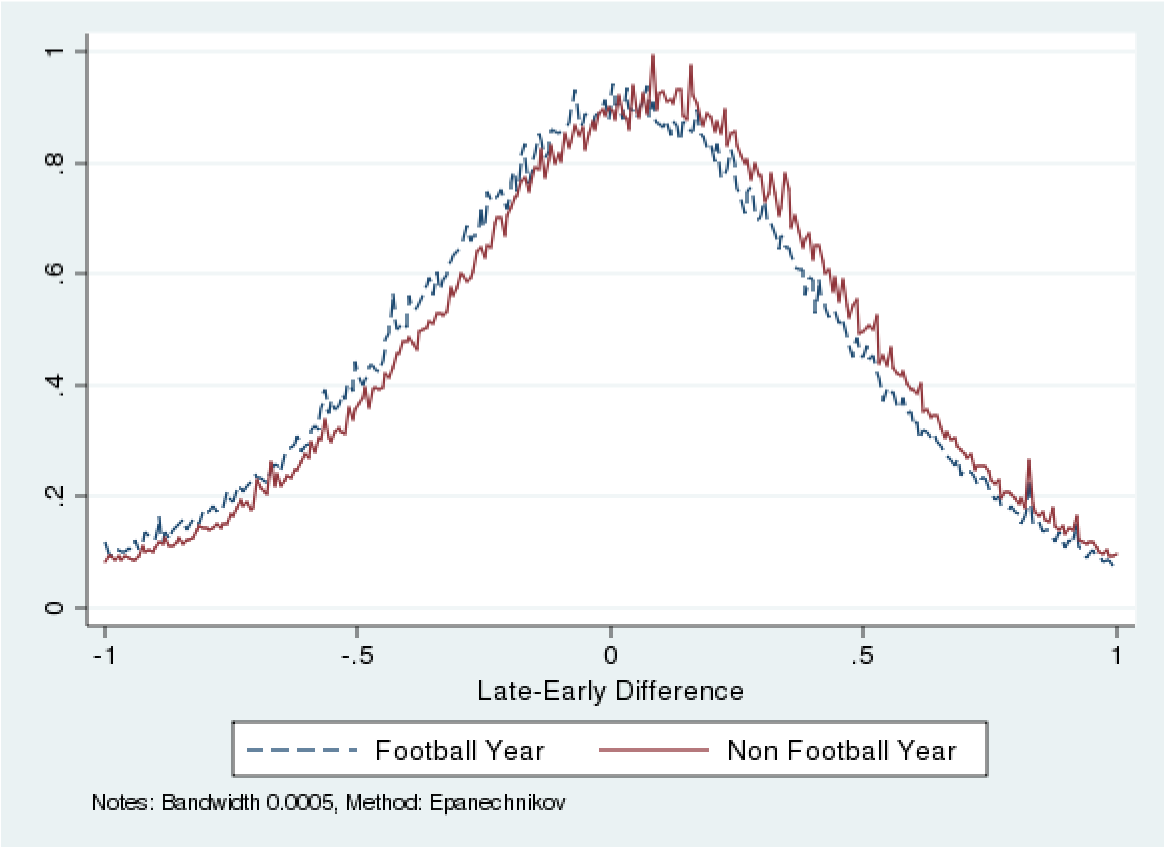
The difference in monthly TV viewing figures, June minus April.




## Figure 2: Annual percentage change in the percentage of pupils obtaining five good GCSEs

Tournament years highlighted.

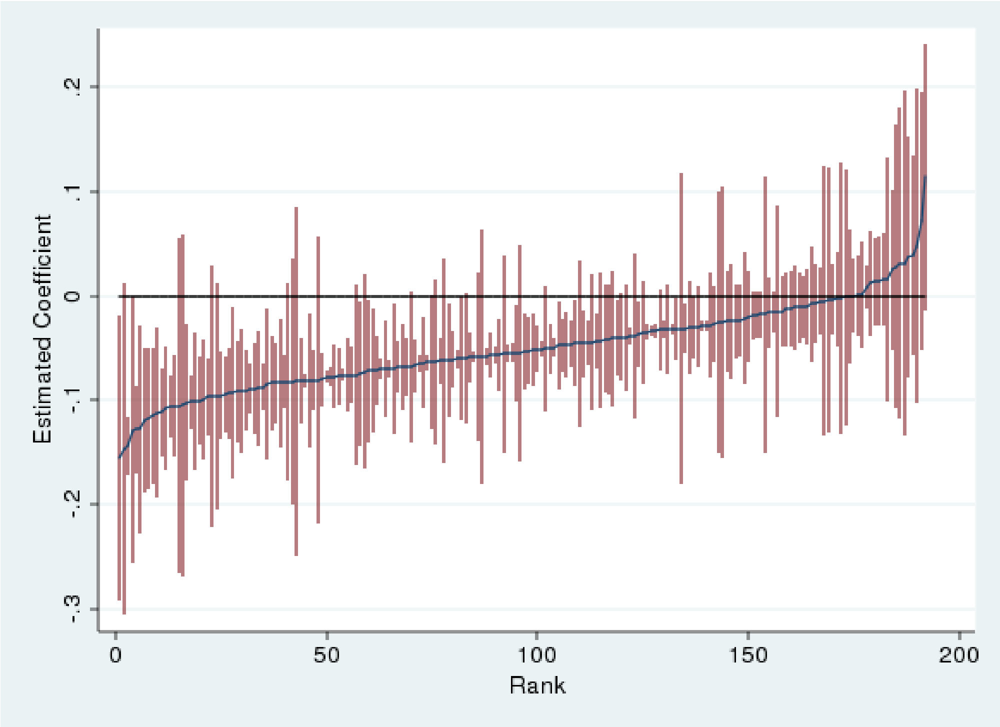
## Figure 3: Density functions for (late-early) subject score difference



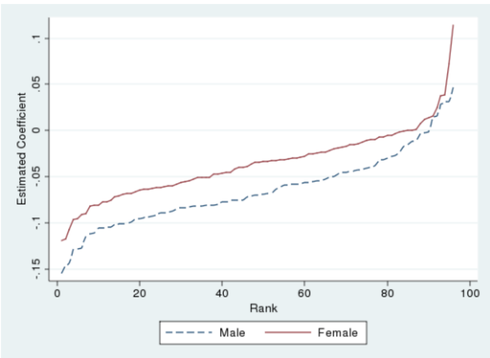
## Figure 4: Difference-in-differences by matched groups

Metric is subject-level SD units.

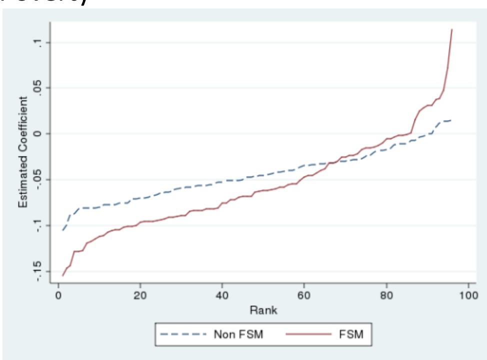
Standard error bars are shown in the main panel, omitted from the group ones for clarity.



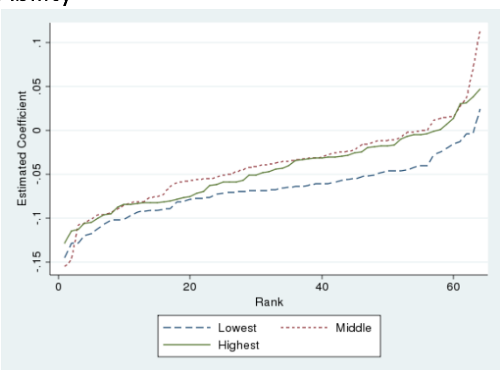

Gender



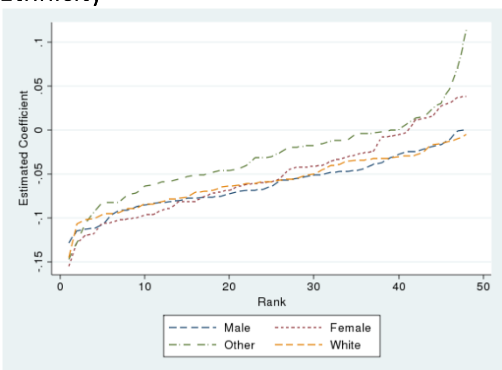

Poverty



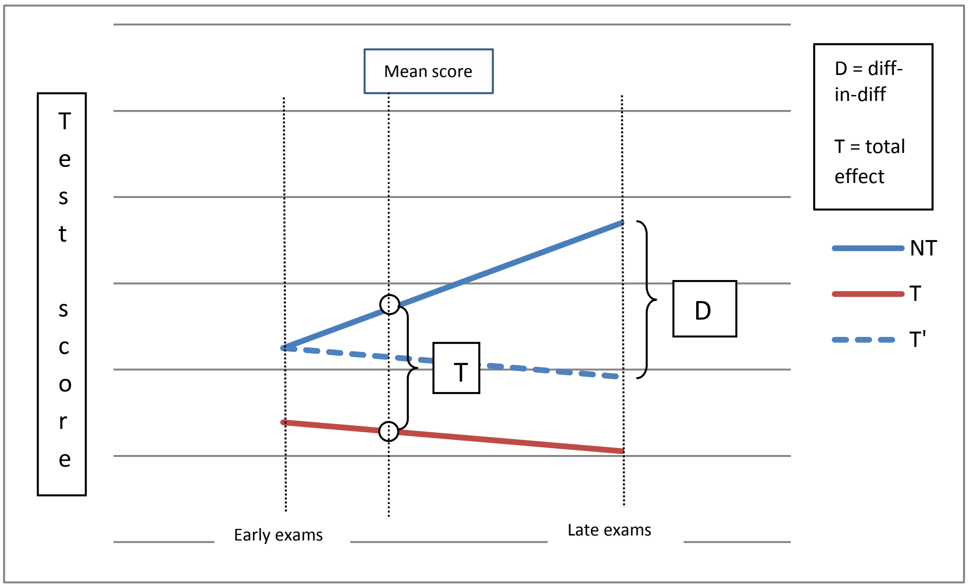

Ability




Ethnicity



## Figure 5: Comparing the difference in difference and the total effect



## Table 1: Data Descriptives

| | All | With both "late" and "early" subjects |
|---|---|---|
| | % | % |
| Male | 50.15 | 49.27 |
| FSM Eligible | 12.05 | 11.03 |
| SEN – non-statemented | 13.48 | 11.40 |
| SEN – statemented | 2.03 | 1.53 |
| Selected ethnicities* | | |
| White | 84.64 | 84.05 |
| Black Caribbean | 1.34 | 1.38 |
| Indian | 2.33 | 2.47 |
| Pakistani | 2.28 | 2.37 |
| GCSE score, normalised | -0.041 | 0.014 |
| Keystage 2 score | 27.03 | 27.34 |
| Number of students | 3,651,667 | 2,970,694 |
| Total observations (subjects*students) | 25,705,081 | 21,963,321 |

Notes: Seven years of data, 2002 – 2008.
One cohort of students per year; final year cohort.
* Full set used in regressions.

## Table 2: Simple Average Differences

(Mean GCSE scores in football-years) – (Mean GCSE scores in non-football-years)
Metric is SD of student average score

| Prior Attainment | | Not Eligible for FSM | | Eligible for FSM | | All pupils |
|---|---|---|---|---|---|---|
| | | Female | Male | Female | Male | |
| Lowest | *Coeff* | 0.0508*** | 0.0357*** | 0.0210*** | 0.0146** | 0.0369*** |
| | | (0.0030) | (0.0030) | (0.0058) | (0.0061) | (0.0024) |
| | *N* | 482,567 | 541,405 | 124,231 | 122,623 | 1,270,826 |
| Middle | *Coeff* | 0.0312*** | 0.0151*** | -0.0012 | -0.0105 | 0.0211*** |
| | | (0.0023) | (0.0025) | (0.0070) | (0.0076) | (0.0020) |
| | *N* | 562,183 | 549,089 | 68,469 | 61,583 | 1,241,324 |
| Highest | *Coeff* | -0.0206*** | -0.0487*** | -0.0890*** | -0.1225*** | -0.0419*** |
| | | (0.0022) | (0.0025) | (0.0094) | (0.0103) | (0.0019) |
| | *N* | 550,510 | 525,954 | 32,222 | 30,831 | 1,139,517 |
| All Pupils | *Coeff* | 0.0133*** | -0.0026 | -0.0178*** | -0.0258*** | -0.0014 |
| | | (0.0021) | (0.0022) | (0.0051) | (0.0054) | (0.0018) |
| | *N* | 1,595,260 | 1,616,448 | 224,922 | 215,037 | 3,651,667 |

Notes: Standard errors are clustered by school.
* indicates significant at 10%, ** indicates significant at 5%, *** indicates significant at 1%.
Prior attainment groups are approximate thirds of the distribution. The boundaries of the groups are adjusted slightly through time as KS2 marks have trended upwards.
The GCSE data are detrended by subtracting a common linear time trend.

## Table 3: Regression analysis of (late – early) difference

| Unit is Individual student<br>Metric is subject level SD | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Tournament year | -0.063*** | -0.054*** | -0.063*** | -0.050*** |
| | (0.001) | (0.002) | (0.001) | (0.002) |
| Tournament year interacted with: | | | | |
| Male | | -0.026*** | | -0.025*** |
| | | (0.002) | | (0.002) |
| FSM | | -0.021*** | | -0.019*** |
| | | (0.002) | | (0.002) |
| Low prior attainment | | -0.009*** | | -0.009*** |
| | | (0.002) | | (0.002) |
| High prior attainment | | -0.011*** | | -0.011*** |
| | | (0.002) | | (0.002) |
| School Fixed effects | | | Y | Y |
| Observations | 2970694 | 2970694 | 2970694 | 2970694 |
| Number of schools | | | 3283 | 3283 |
| R-squared | 0.03 | 0.03 | 0.03 | 0.03 |

Notes: Standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
Dependent variable is the difference between the individual's mean score in late subjects and early subjects
Variables also included as main effects are: pupil gender, ethnicity, FSM status, SEN status, prior attainment, and location, plus year dummies. Only a selection of interactions are shown.

## Table 4: Group difference-in-differences

Metric is subject-level SD

| Prior Attainment | | Not Eligible for FSM | | Eligible for FSM | | All pupils |
|---|---|---|---|---|---|---|
| | | Female | Male | Female | Male | |
| Lowest | *Coeff* | -0.0584*** | -0.0679*** | -0.0649*** | -0.1077*** | -0.0680*** |
| | | (0.0032) | (0.0035) | (0.0057) | (0.0065) | (0.0029) |
| | *N* | 374,455 | 396,640 | 89,888 | 82,763 | 943,746 |
| Middle | *Coeff* | -0.0253*** | -0.0740*** | -0.0208*** | -0.0993*** | -0.0495*** |
| | | (0.0027) | (0.0031) | (0.0060) | (0.0074) | (0.0025) |
| | *N* | 471,834 | 445,225 | 54,480 | 46,750 | 1,018,289 |
| Highest | *Coeff* | -0.0343*** | -0.0661*** | -0.0385*** | -0.0755*** | -0.0507*** |
| | | (0.0027) | (0.0028) | (0.0075) | (0.0082) | (0.0023) |
| | *N* | 488,550 | 466,204 | 27,816 | 26,089 | 1,008,659 |
| All Pupils | *Coeff* | -0.0385*** | -0.0680*** | -0.0471*** | -0.0991*** | -0.0556*** |
| | | (0.0022) | (0.0025) | (0.0043) | (0.0050) | (0.0021) |
| | *N* | 1,334,839 | 1,308,069 | 172,184 | 155,602 | 2,970,694 |

Notes: Standard errors are clustered by school
* indicates significant at 10%, ** indicates significant at 5%, *** indicates significant at 1%.
Prior attainment groups are approximate thirds of the distribution. The boundaries of the groups are adjusted slightly through time as KS2 marks have trended upwards.

## Table 5: Quantiles of differences-in-differences for matched school-groups

Metric is subject-level SD

| | {(Late – early) in tournament} - {(Late – early) in non- tournament} | | | | | | |
|---|---|---|---|---|---|---|---|
| | **p5** | **p10** | **p25** | **p50** | **p75** | **p90** | **p95** |
| All Pupils | -0.3307 | -0.2570 | -0.1531 | -0.0486 | 0.0577 | 0.1571 | 0.2150 |
| Male | -0.3571 | -0.2846 | -0.1756 | -0.0628 | 0.0489 | 0.1531 | 0.2134 |
| FSM | -0.4215 | -0.3556 | -0.2247 | -0.0339 | 0.1017 | 0.2251 | 0.2546 |
| Low ability | -0.4006 | -0.3170 | -0.1854 | -0.0495 | 0.0814 | 0.1965 | 0.2632 |
| Middle ability | -0.3380 | -0.2711 | -0.1579 | -0.0457 | 0.0731 | 0.1792 | 0.2407 |
| High ability | -0.2987 | -0.2374 | -0.1444 | -0.0491 | 0.0477 | 0.1364 | 0.1912 |

Notes: These figures are based on school-group matching. For all pupils, there are 14,940 school-groups. School-groups are only included if there are at least 20 students within the school-group in both tournament and non-tournament years. Quantiles of the distribution of the following statistic are reported:

$$\Delta_{sg} = \overline{(\bar{y}_{isg,late} - \bar{y}_{isg,early})}_{tournament} - \overline{(\bar{y}_{isg,late} - \bar{y}_{isg,early})}_{non-tournament}$$

## Table 6: Student*subject fixed effect regression results

Observation = student*subject
Metric is subject-level SD
Observation = student*subject
Metric is subject-level SD

| | (1) | (2) | (3) |
|---|---|---|---|
| Proportion of exams within subject which are "late" | 0.068*** | 0.103*** | 0.126*** |
| | (0.001) | (0.001) | (0.001) |
| Proportion of exams within subject which are "late" * Year is a tournament year | -0.007*** | -0.009*** | -0.009*** |
| | (0.001) | (0.001) | (0.001) |
| Year dummies | Y | Y | |
| Student Characteristics | | Y | |
| Student fixed effects | | | Y |
| Number of observations | 25,705,081 | 25,705,081 | 25,705,081 |
| Number of pupils | 3,651,667 | 3,651,667 | 3,651,667 |

Notes: 1. Standard errors in parentheses; standard errors clustered at student level.
2. * significant at 10%; ** significant at 5%; *** significant at 1%
3. "Late" is defined by calendar date for all years, and coincides with the tournament dates; see text for details
4. Student characteristics are: gender, ethnicity, month of birth, poverty status, SEN status, English as additional language, prior ability measures (Keystage 2 English score, Keystage 2 maths score, Keystage 2 Science score)

## Table 7: Quantifying the Results: Other Metrics

| | Impact of effort on exams | Overall Effect Metric: SD of pupil mean score | Overall Effect Metric: GCSE grades |
|---|---|---|---|
| **Difference in difference** | | | |
| Table 3 | | | |
| Mean (col. 3) | -0.126 | -0.015 | -0.208 |
| Poor, male, white, low attainment (col. 4) | -0.206 | -0.025 | -0.347 |
| Table 5 | | | |
| All pupils, (median) | -0.116 | -0.014 | -0.194 |
| All pupils, (p10) | -0.202 | -0.025 | -0.347 |
| Male pupils, (median) | -0.140 | -0.017 | -0.236 |
| Male pupils, (p10) | -0.216 | -0.026 | -0.361 |

Notes: Column 1 = coefficient * 2
Column 2 = column 1* (1.76/7.80)*(11.54/10.68) {share of late exams}*{converting subject sd to pupil sd}
Column 3 = column 2*(10.68/6)*7.80 {converting to gcse points} {converting to letter grades} multiplying by the number of exams.

# Appendix Tables

**Appendix Table 1: The football tournaments from 2002 to 2008**

| Year | Host country | Tournament | Did England qualify? | Start date | End date |
|---|---|---|---|---|---|
| 2002 | South Korea and Japan | World Cup | Yes | 31st May | 30th June |
| 2004 | Portugal | European championships | Yes | 12th June | 4th July |
| 2006 | Germany | World Cup | Yes | 9th June | 9th July |
| 2008 | Austria and Switzerland | European championships | No | 7th June | 29th June |

**Appendix Table 2: The examination dates from 2002 to 2008**

| Year | 'Football' year | Examination start date | Examination end date | % of exams during football |
|---|---|---|---|---|
| 2002 | Yes | 13th May | 28th June | 61% |
| 2003 | No | 12th May | 27th June | - |
| 2004 | Yes | 17th May | 30th June | 49% |
| 2005 | No | 16th May | 30th June | - |
| 2006 | Yes | 15th May | 28th June | 48% |
| 2007 | No | 14th May | 27th June | - |
| 2008 | Yes | 13th May | 25th June | 46% |

**Appendix Table 3: Subjects that do and do not have overlap with the football tournaments in 2002, 2004 and 2006**

| 2002<br><br>Subjects that have no exams during the football | 2002<br><br>Subjects that have some exams during the football | 2002<br><br>Subjects that have all their exams during the football |
|---|---|---|
| D&T automobile studies | Biology | Economics |
| Drama | Business Studies | Greek |
| English Lit | Chemistry | History |
| Geography | Classical civilisation | Home economics |
| German | D&T Electronics | Information Technology |
| Health studies | D&T Food | Physics |
| Humanities | D&T Graphics | Psychology |
| Information Studies | D&T Industrial | Social Science |
| Music | D&T Resistant Materials | Sociology |
| Persian | D&T Systems and Control Technology | |
| Physical Education | D&T Textiles Technology | |
| | French | |
| | English Language | |
| | Latin | |
| | Mathematics | |
| | Portuguese | |
| | Religious Studies | |
| | Religious Education | |
| | Science (both double and single) | |
| | Turkish | |

| 2004<br><br>Subjects that have no exams during the football | 2004<br><br>Subjects that have some exams during the football | 2004<br><br>Subjects that have all their exams during the football |
|---|---|---|
| Applied ICT | Biblical Hebrew | Applied Business |
| Applied Science | Biology | Chemistry |
| Business and communication systems | Business Studies | Classical Civilisation |
| Citizenship studies | D&T Electronics | Classical Greek |
| D&T Systems and Control technology | D&T Food | Economics |
| Drama | D&T Graphics | History |
| Engineering | D&T Industrial | Home economics |
| Geography | D&T Resistant Materials | Leisure and Tourism |
| German | D&T Textiles Technology | Media Studies |
| Health and Social Care | English Language | Physics |
| Humanities | English Lit | Psychology |
| ICT | French | Sociology |
| Latin | Mathematics | |
| Manufacturing | Religious Studies | |
| Music | Science (both double and single) | |
| Persian | | |
| Physical Education | | |
| Portuguese | | |
| Rural and Agricultural Science | | |
| Spanish | | |
| Turkish | | |

| 2006 | 2006 | 2006 |
|---|---|---|
| **Subjects that have no exams during the football** | **Subjects that have some exams during the football** | **Subjects that have all their exams during the football** |
| Applied ICT | Biblical Hebrew | Applied Business |
| Applied Science | Biology | Chemistry |
| Business and communication systems | Business Studies | Classical Civilisation |
| Citizenship studies | D&T Electronics | Classical Greek |
| D&T Systems and Control technology | D&T Food | Economics |
| Drama | D&T Graphics | History |
| Engineering | D&T Industrial | Home economics |
| Geography | D&T Resistant Materials | Leisure and Tourism |
| German | D&T Textiles Technology | Media Studies |
| Health and Social Care | English Language | Physics |
| Humanities | English Lit | Psychology |
| ICT | French | Sociology |
| Latin | Mathematics | |
| Manufacturing | Religious Studies | |
| Music | Science (both double and single) | |
| Persian | | |
| Physical Education | | |
| Portuguese | | |
| Rural and Agricultural Science | | |
| Spanish | | |
| Turkish | | |

# Appendix Table 4: Football programmes in the top 10 most viewed programmes

| Channel and programme rank for that week | | Week ending and programme | Viewers (millions) |
|---|---|---|---|
| **2002** | | | |
| BBC1 | | w/e 9th June 2002 | |
| | 2 | WORLD CUP 2002: ARGENTINA V ENGLAND (FRI 1230) | 12 |
| | 5 | WORLD CUP 2002: POST-MATCH (FRI 1420) | 10.49 |
| BBC1 | | w/e 16th June 2002 | |
| | 2 | WORLD CUP 2002: ENGLAND V DENMARK (SAT 1230) | 12.47 |
| | 4 | WORLD CUP 2002: ENGLAND V NIGERIA (WED 0730) | 12.22 |
| | 7 | WORLD CUP 2002: POSTMATCH (SAT 1420) | 8.85 |
| | 9 | WORLD CUP 2002: SPAIN V IRELAND (SUN 1230) | 7.77 |
| BBC1 | | w/e 23rd June 2002 | |
| | 1 | WORLD CUP 2002: ENGLAND V BRAZIL (FRI 0730) | 12.46 |
| | 6 | WORLD CUP 2002: POST-MATCH (FRI 0920) | 9.77 |
| BBC1 | | w/e 30th June 2002 | |
| | 4 | WORLD CUP 2002: GERMANY V BRAZIL (SUN 1200) | 10.08 |
| | 7 | WORLD CUP 2002: POST MATCH (SUN 1350) | 8.95 |
| **2004** | | | |
| BBC1 | | w/e 13th June 2004 | |
| | 7 | EURO 2004: SPA V RUS (SAT 1945) | 6.4 |
| | 8 | EURO 2004: PORT V GRC (SAT 1700) | 6.19 |
| ITV1 | | 13th June 2004 | |
| | 1 | EURO 2004 FRA V ENG (SUN 1944) | 17.8 |
| BBC1 | | w/e 20th June 2004 | |
| | 4 | EURO 2004: SPA V PORT (SUN 1945) | 8.78 |
| | 5 | EURO 2004: GER V NETH (TUE 1945) | 7.95 |
| | 6 | EURO 2004: CRO V FRA (THU 1946) | 7.55 |
| | 7 | EURO 2004: POST-MATCH (THU 2135) | 7.23 |
| | 8 | EURO 2004: POST-MATCH (SUN 2135) | 6.85 |
| | 9 | EURO 2004: NETH V CZECH (SAT 1945) | 6.74 |
| ITV1 | | w/e 20th June 2004 | |

| | | | |
|---|---|---|---|
| | 1 | EURO 2004 ENG V SWI (THU 1659) | 14.31 |
| BBC1 | | w/e 27th June 2004 | |
| | 1 | EURO 2004: POR V ENG (THU 1945) | 20.66 |
| | 2 | EURO 2004: CRO V ENG (MON 1945) | 18.28 |
| | 3 | EURO 2004: POST-MATCH (MON 2136) | 14.48 |
| | 4 | EURO 2004: POST-MATCH (THU 2229) | 14.22 |
| | 5 | EURO 2004: PREMATCH (THU 1929) | 11.71 |
| | 7 | EURO 2004: PREMATCH (MON 1929) | 9.83 |
| ITV1 | | w/e 27th June 2004 | |
| | 4 | EURO 2004 GER V CZE (WED 1944) | 8.28 |
| | 9 | EURO 2004 SWE V NETH (SAT 1945) | 7.04 |
| **2006** | | | |
| BBC1 | | w/e 4th June 2006 | |
| | 1 | MATCH OF THE DAY LIVE (TUE 1958) | 9.29 |
| BBC1 | | w/e 11th June 2006 | |
| | 1 | WORLD CUP 2006: ENG V PAR (SAT 1400) | 12 |
| | 2 | WORLD CUP 2006: POST-MATCH (SAT 1551) | 9.29 |
| | 10 | WORLD CUP 2006: GER V CRI (FRI 1701) | 5.65 |
| BBC1 | | w/e 18th June 2006 | |
| | 1 | WORLD CUP 2006: BRA V CRO (TUE 2000) | 9.64 |
| | 2 | WORLD CUP 2006: GER V POL (WED 2000) | 8.11 |
| | 4 | WORLD CUP 2006: POST-MATCH (WED 2149) | 6.74 |
| | 5 | WORLD CUP 2006: ITA V GHA (MON 2000) | 6.69 |
| | 8 | WORLD CUP 2006: POST-MATCH (SUN 2151) | 6.39 |
| | 9 | WORLD CUP 2006: POST-MATCH (TUE 2151) | 6.38 |
| | 10 | WORLD CUP 2006: FRA V KOR (SUN 2000) | 6.17 |
| ITV1 | | w/e 18th June 2006 | |
| | 1 | WORLD CUP 06: ENG V TRI (THU 1650) | 13.67 |
| | 5 | WORLD CUP 06: BRA V AUS (SUN 1658) | 8.08 |
| | 10 | WORLD CUP 06: SWE V PAR (THU 1959) | 6.63 |
| BBC1 | | w/e 25th June 2006 | |
| | 1 | WORLD CUP 2006: ENG V ECU (SUN 1600) | 16.29 |
| | 2 | WORLD CUP 2006: POST-MATCH (SUN 1750) | 13.45 |
| | 3 | WORLD CUP 2006: ARG V MEX (SAT 2000) | 8.46 |
| | 4 | WORLD CUP 2006: JAP V BRA (THU 2000) | 7.81 |
| | 10 | WORLD CUP 2006: PREMATCH (SUN 1529) | 7.44 |
| ITV1 | | w/e 25th June 2006 | |
| | 1 | WORLD CUP 06 (TUE 1950) | 18.46 |
| | 3 | WORLD CUP 06 (WED 1958) | 8.74 |
| | 7 | WORLD CUP 06 (SUN 1958) | 7.43 |
| | 9 | WORLD CUP 06: PREMATCH (TUE 1903) | 6.7 |
| **2008** | | | |
| BBC1 | | w/e 15th June 2008 | |
| | 10 | EURO 2008: MATCH OF THE DAY LIVE (FRI 1929) | 5.58 |
| ITV1 | | w/e 15th June 2008 | |
| | 10 | EURO 2008 LIVE (MON 1929) | 5.74 |
| BBC1 | | w/e 22nd June 2008 | |
| | 4 | EURO 2008: MATCH OF THE DAY LIVE (SUN 1930) | 7.21 |
| | 5 | EURO 2008: MATCH OF THE DAY LIVE (TUE 1929) | 6.29 |
| | 7 | EURO 2008: MATCH OF THE DAY LIVE (FRI 1929) | 5.64 |
| ITV1 | | w/e 22nd June 2008 | |
| | 3 | EURO 2008 LIVE (SAT 1929) | 7.37 |
| | 5 | EURO 2008 LIVE (THU 1929) | 6.89 |
| BBC1 | | w/e 29th June 2008 | |
| | 1 | EURO 2008: MATCH OF THE DAY LIVE (SUN 1856) | 8.84 |
| | 6 | EURO 2008: MATCH OF THE DAY LIVE (WED 1929) | 6.95 |
| ITV1 | | w/e 29th June 2008 | |
| | 6 | EURO 2008 LIVE (THU 1929) | 6.77 |